doi:10.1016/j.jmb.2005.12.062 J. Mol. Biol. (2006) 357, 82–99

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Structural Mimicry of CD4 by a Cross-reactive HIV-1 Neutralizing Antibody with CDR-H2 and H3 Containing Unique Motifs

Ponraj Prabakaran[1], Jianhua Gan[2], You-Qiang Wu[1], Mei-Yun Zhang[1,3] Dimiter S. Dimitrov[1]* and Xinhua Ji[2]*

[1]*Protein Interactions Group Center for Cancer Research Nanobiology Program, National Cancer Institute, NIH, Frederick MD 21702, USA*

[2]*Biomolecular Structure Section, Macromolecular Crystallography Laboratory Center for Cancer Research National Cancer Institute NIH, Frederick, MD 21702 USA*

[3]*Basic Research Program SAIC-Frederick, Inc., Frederick MD 21702, USA*

**Corresponding author*

Human immunodeficiency virus (HIV) entry into cells is initiated by the binding of its envelope glycoprotein (Env) gp120 to receptor CD4. Antibodies that bind to epitopes overlapping the CD4-binding site (CD4bs) on gp120 can prevent HIV entry by competing with cell-associated CD4; their ability to outcompete CD4 is a major determinant of their neutralizing potency and is proportional to their avidity. The breadth of neutralization and the likelihood of the emergence of antibody-resistant virus are critically dependent on the structure of their epitopes. Because CD4bs is highly conserved, it is reasonable to hypothesize that antibodies closely mimicking CD4 could exhibit relatively broad cross-reactivity and a high probability of preventing the emergence of resistant viruses. Previously, in a search for antibodies that mimic CD4 or the co-receptor, we identified and characterized a broadly cross-reactive HIV-neutralizing CD4bs human monoclonal antibody (hmAb), m18. Here, we describe the crystal structure of Fab m18 at 2.03 Å resolution, which reveals unique conformations of heavy chain complementarity-determining regions (CDRs) 2 and 3 (H2 and H3). H2 is highly bulged and lacks cross-linking interstrand hydrogen bonds observed in all four canonical structures. H3 is 17.5 Å long and rigid, forming an extended β-sheet decorated with an α-turn motif bearing a phenylalanine-isoleucine fork at the apex. It shows striking similarity to the Ig CDR2-like C′C″ region of the CD4 first domain D1 that dominates the binding of CD4 to gp120. Docking simulations suggest significant similarity between the m18 epitope and the CD4bs on gp120. Fab m18 does not enhance binding of CD4-induced (CD4i) antibodies, nor does it induce CD4-independent fusion mediated by the HIV Env. Thus, vaccine immunogens based on the m18 epitope structure are unlikely to elicit antibodies that could enhance infection. The structure can also serve as a basis for the design of novel, highly efficient inhibitors of HIV entry.

Published by Elsevier Ltd.

Keywords: HIV; antibody; crystal structure; gp120; vaccine

Introduction

Virus entry into animal cells is initiated by binding to cell surface-associated receptors.[1,2] The human immunodeficiency virus (HIV) envelope glycoprotein (Env) gp120 binds to receptor CD4, triggering a cascade of events leading to virus–cell fusion and infection. Antibodies elicited by natural infection or immunization can interfere with this process by binding to Env and interfering with certain stages of the viral entry process. However, HIV has evolved to resist the neutralizing activity of antibodies by various strategies, including rapid generation of mutants with altered structures of antibody epitopes. During the long chronic infection phase, the immune system

Abbreviations used: CD4bs, CD4-binding site; CD4i, CD4-induced; CDR, complementarity-determining region; hmAb, human monoclonal antibody; HIV, human immunodeficiency virus; SIV, simian immunodeficiency virus.

E-mail addresses of the corresponding authors: dimitrov@ncifcrf.gov; jix@ncifcrf.gov

0022-2836/$ - see front matter Published by Elsevier Ltd.

responds by generating a variety of antibodies, some of which exhibit unique properties. Of the many known human monoclonal antibodies (hmAbs) that bind the Env, only a few exhibit potent and broad HIV-1 neutralizing activity *in vitro* and can prevent HIV-1 infection in animal models.[3–5] These antibodies target conserved viral structures that are critical for the mechanism of cell entry. The CD4-binding site (CD4bs) on gp120 is highly conserved and is an obvious target for such antibodies.

Many hmAbs with epitopes that overlap the CD4bs on gp120 have been characterized, including b12,[6–12] b6,[9,12] 15e,[13] 5145A,[14] F105,[15] F91,[16] 1125H,[17] 21h,[13,18] 654-30D,[19] m14,[20] and m18.[21] The molecular mechanisms of antibody function, which determine the potency and breadth of HIV-1 neutralization, have been studied most extensively for b12.[11,22] Although CD4bs antibodies frequently exhibit broad reactivity with monomeric gp120, b12 is unique in neutralizing many HIV-1 isolates from different clades.[9,23,24] The difference in the neutralizing activity between b12 and most CD4bs antibodies was ascribed to the ability of b12 to bind with high affinity to both monomeric and trimeric forms of gp120, whereas most CD4bs antibodies bind predominantly to the monomeric form. Another unique feature of b12 related to its potent HIV-neutralizing activity is the lack of significant conformational changes upon its interaction with gp120, in contrast to other anti-gp120 antibodies.[25] The crystal structure of b12 revealed a long, protruding, and rigid complementarity-determining region (CDR) H3, which could reach deep inside the CD4bs on gp120.[22] Long H3s have been observed also in CD4-induced (CD4i) antibodies (antibodies targeting hidden epitopes on gp120, which become highly exposed after CD4 binding), such as Fab 17b[26,27] and Fab X5,[28,29] and anti-gp41 antibodies, such as 2F5,[30,31] and 4E10.[32] It appears that these antibodies contain long, protruding H3s that can reach recessed binding sites. Interestingly, some CD4i antibodies (e.g. 412d) mimic certain aspects of the HIV co-receptor CCR5, including tyrosine sulfation of its N terminus.[27,33,34]

In a search for antibodies that could mimic CD4 or co-receptor, we have recently identified a cross-reactive HIV hmAb Fab m18 that binds to a variety of recombinant soluble Envs from different clades, and can inhibit cell fusion and virus entry mediated by Envs from primary HIV isolates.[21] Here, we present the crystal structure of Fab m18 and compare the structure with the CD4bs antibodies b12[22] and F105.[35] The m18 structure reveals unique H2 and H3 conformations. The H3 motif is rigid and protruding; it strikingly resembles the CDR2-like loop C′C″ of the CD4 first domain D1, which dominates the binding of CD4 to gp120 as observed in the crystal structure of the CD4–gp120–17b complex.[26] Docking simulations of the m18–gp120 complex using the crystal structures of CD4-bound[26] and CD4-free[36] gp120, along with mutagenesis data, suggest similar interactions of m18 and CD4 with gp120.

Results and Discussion

Overall structure of Fab m18

Although many CD4bs antibodies have been characterized,[11,18] structural information is available only for b12.[22] The sequences of the variable domains of the heavy (V_H) and the light (V_L) chains of Fab m18 are 46% identical with and 63% similar to those of b12, suggesting a significant degree of similarity between the frameworks supporting the combining sites of the two antibodies. However, their hypervariable regions are significantly different. The sequences of the constant domains of the heavy (C_{H1}) and light (C_L) chains of Fab m18 are identical with those of b12, except for two mutations in each chain. Structural alignment between m18 and b12 based on the constant domains revealed a notable shifting of the variable domains between the two antibodies. This difference in the orientation of variable domains between the two structures caused the failure of our molecular replacement attempt using the entire Fab b12 as the search model. Using the constant domains of b12 alone, we obtained outstanding molecular replacement solutions. The locations of variable domains were revealed by difference Fourier synthesis as described in Materials and Methods. The data collection parameters and refinement statistics are given in Table 1. The crystal structure of m18 (Figure 1(a)) contains three Fab molecules in the asymmetric unit, which are referred to as Mol A (chains A and B), Mol B (chains C and D) and Mol C (chains E and F),

Table 1. Data collection parameters and refinement statistics for Fab m18

Space group	$P2_1$
Unit cell parameters	
a (Å)	48.6
b (Å)	82.6
c (Å)	187.2
β (deg.)	95.3
Resolution range (Å)	29.25–2.03
No. unique reflections	81,882
Completeness (%)	85.8 (69.1)[a]
$I/\sigma(I)$	3.5 (7.6)[a]
R_{merge} (%)	0.087 (0.316)[a]
R-factor (%)	22.5
R_{free} (%)	26. 9
No. amino acid residues, average *B*-factor (Å^2)	1307, 31.7
No. water molecules, average *B*-factor (Å^2)	499, 31.4
RMSD bond lengths (Å)	0.007
RMSD bond angles (deg.)	1.4
Ramachandran plot	
Most favored φ and ψ angles (%)	86.5
Disallowed φ and ψ angles (%)	0.6

[a] Values in parentheses are values for the outmost shell of reflections, 2.10–2.03 Å.

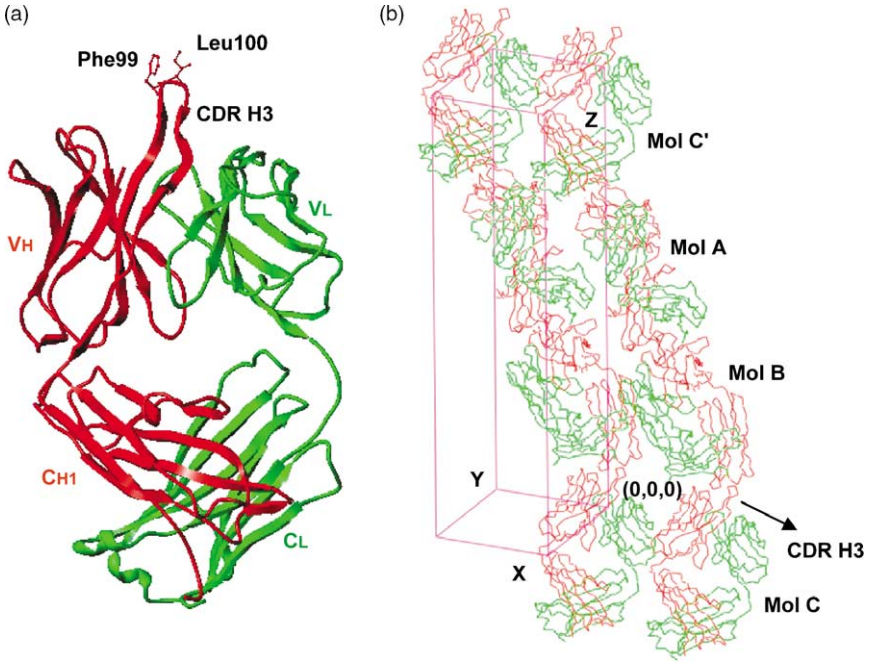

Figure 1. Overall structure of Fab m18. (a) The ribbon diagram shows one of the three m18 molecules in the asymmetric unit; the light and heavy chains are in green and red, respectively. The phenylalanine and leucine residues at the tip of the H3 are indicated. (b) A packing diagram shows the arrangement of Fab m18 molecules in the unit cell. H3, which is indicated with an arrow, facilitates Fab–Fab interactions in the crystal lattice.

respectively. Chains A, C, and E are the light chains, and B, D, F are the heavy chains. The three Fab molecules are packed tightly in a column along the crystallographic 2_1 axis, interacting with each other in a head-to-tail fashion, with the V_H domain of one Fab contacting the C_{H1} domain of another (Figure 1(b)). Specifically, H3 mediates the Fab–Fab interactions by inserting the side-chain of Phe99 into a hydrophobic pocket in the C_{H1} surface of the neighboring Fab molecule, burying a surface area of 146 Å^2. This closely resembles the gp120–CD4 interaction in which the hotspot residue Phe43 of CD4 is inserted into the second hydrophobic pocket of the gp120 surface, resulting in a buried surface area of 152 Å^2.[26] The averaged temperature factors of the three Fab molecules are within the range of 30–32 Å^2 and the root-mean-square (RMS) deviations for C^α positions range from 0.4–0.8 Å among the three molecules. The buried surface areas between Mol A and Mol B, Mol B and Mol C, and Mol C′ (symmetry-related) and Mol A are 908 Å^2, 997 Å^2, and 1110 Å^2, respectively, indicating some differences in the packing and arrangement of interfacial residues. The electron density maps for Mol A, which define the conformations of H2 and H3 unambiguously, are shown in Figure 2(a) and (b), respectively. The H2 motif contains residues 50–65, among which Tyr52 points to the interior of the motif and makes a hydrogen bond with Asn58. The H3 motif contains residues 95–102 with an insertion of six residues after position 100 and contains hydrophobic residues Phe99 and Leu100, which are exposed at the tip. The Ramachandran plot[37] shows that the φ/ψ torsion angle distributions are similar to those observed in other well-refined structures, including two exceptions (Ala51 and Gln29 in the CDR L2 region) generally observed in many immunoglobulin structures,[38–40] and Ser15 in the framework of one of the heavy chains F. All the outliers have well-defined electron density. Not observed are the N-terminal residue of each chain, three C-terminal residues of chain B, and five C-terminal residues of chains D and E. Also observed are 499 water molecules and three sulfate ions from the crystallization buffer system.

Recently, the crystal structure of a broadly reactive CD4bs antibody F105 has been reported.[35] The structure of F105 also has an extended H3 with a phenylalanine residue at position 100A. The sequences and structures of CD4bs antibodies b12, m18, and F105 are shown in Figure 3. Although the framework of the variable domains displays sequence similarities (Figure 3(a)), the H3s of these antibodies exhibit distinct conformations (Figure 3(b)). The conformational differences of H3s and the arrangement of CD4bs residues of gp120 might cause variations in binding affinities and neutralizing activities.

Domain interactions and packing of interfacial residues

The antibody-combining site is formed as the result of association between V_H and V_L domains, and the amino acids at the interface of the domains are responsible for the relative orientation of hypervariable loops, which determines their shape, properties, and specificity required for antigen binding. The importance of the V_H–V_L association in the diversification of antibody repertoires prompted several investigations to characterize the surfaces involved in their association, the propensities of amino acid residues, and packing interactions at the interface.[41–44] In order to assess the role of the V_H–V_L domain interactions in the m18 structure, we analyzed the interface formed

by the domains and the interaction between the residues at the surface. The total surface areas buried between the V_H and V_L domains are 1770 Å^2, 1730 Å^2, and 1760 Å^2 for molecules A, B, and C, respectively. These values are comparable to the average value of 1570($\pm$160) Å^2 calculated from the analysis of more than 200 Fab structures.[45] Inter-residue contact analysis carried out by applying the distance criterion of being greater than the sum of the van der Waals radii between two atoms, and the packing analysis of side-chains at the interface, were performed by SYBYL7.0 (Tripos Inc., St. Louis, USA). V_L residues L36, L43, L46, L89, L91, L94, and L96, and V_H residues H39, H47, H52, H58, H95, H100C, H100E-F, and H103 form the interface in all three m18 molecules. Most of the residues, including L36, L46, L89, L91, and L96, and H39, H47, H95, and H103, were already found at the interface and are highly conserved among 23 structures of human and murine antibodies.[46] These residues appeared also in the list of 20 residues that were proposed to be involved in the V_H–V_L interaction by using a three-layer packing model.[47] Interestingly, we found that the hypervariable loops in m18 contribute more than 50% of all contacts at the interface.

All six hypervariable CDRs forming the antibody-combining site from the three Fab m18 molecules in the asymmetric unit are superimposed and shown as C^α traces in Figure 4(a). The interactions between L3, H2, and H3 form the major part of the V_H–V_L interface. A unique quadruple tyrosine motif consisting of Tyr94 and Tyr96 from L3, Tyr52 from H2, and Tyr95 at the base of H3 was found at the centre of the interface between the V_H and V_L domains. This hydrogen bonded network of tyrosine residues acts as a cradle and might influence the conformation of H3 indirectly. Additionally, three water molecules are buried at the V_H–V_L interface, and these water molecules interact with framework residues from both the V_H and V_L domains (Figure 4(b)). It appears that water molecules at the V_H–V_L interface play a substantial role in stabilizing the Fab structure. The structure of Fab HyHEL-5 in complex with its antigen hen egg-white lysozyme also has three water molecules trapped at the interface

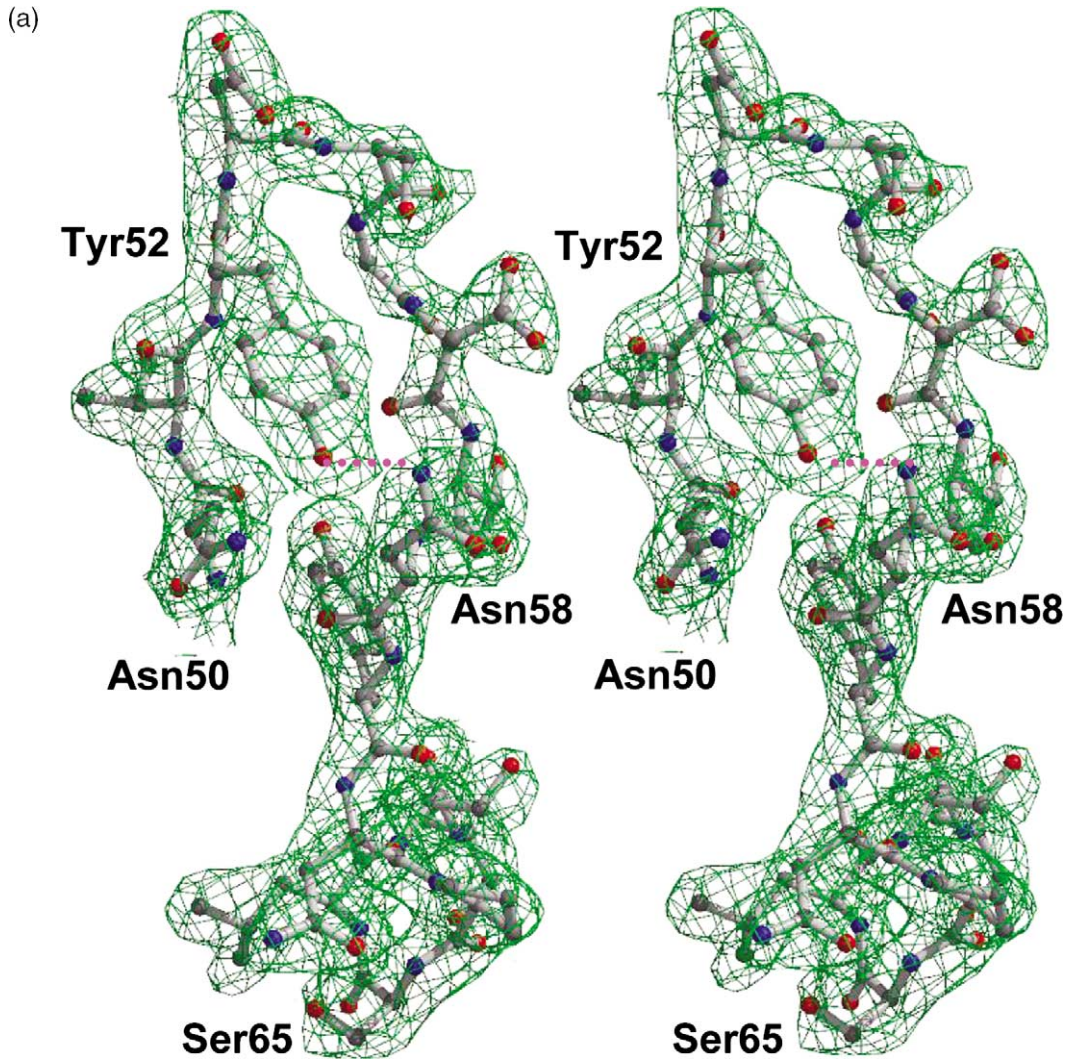

Figure 2 (*legend next page*)

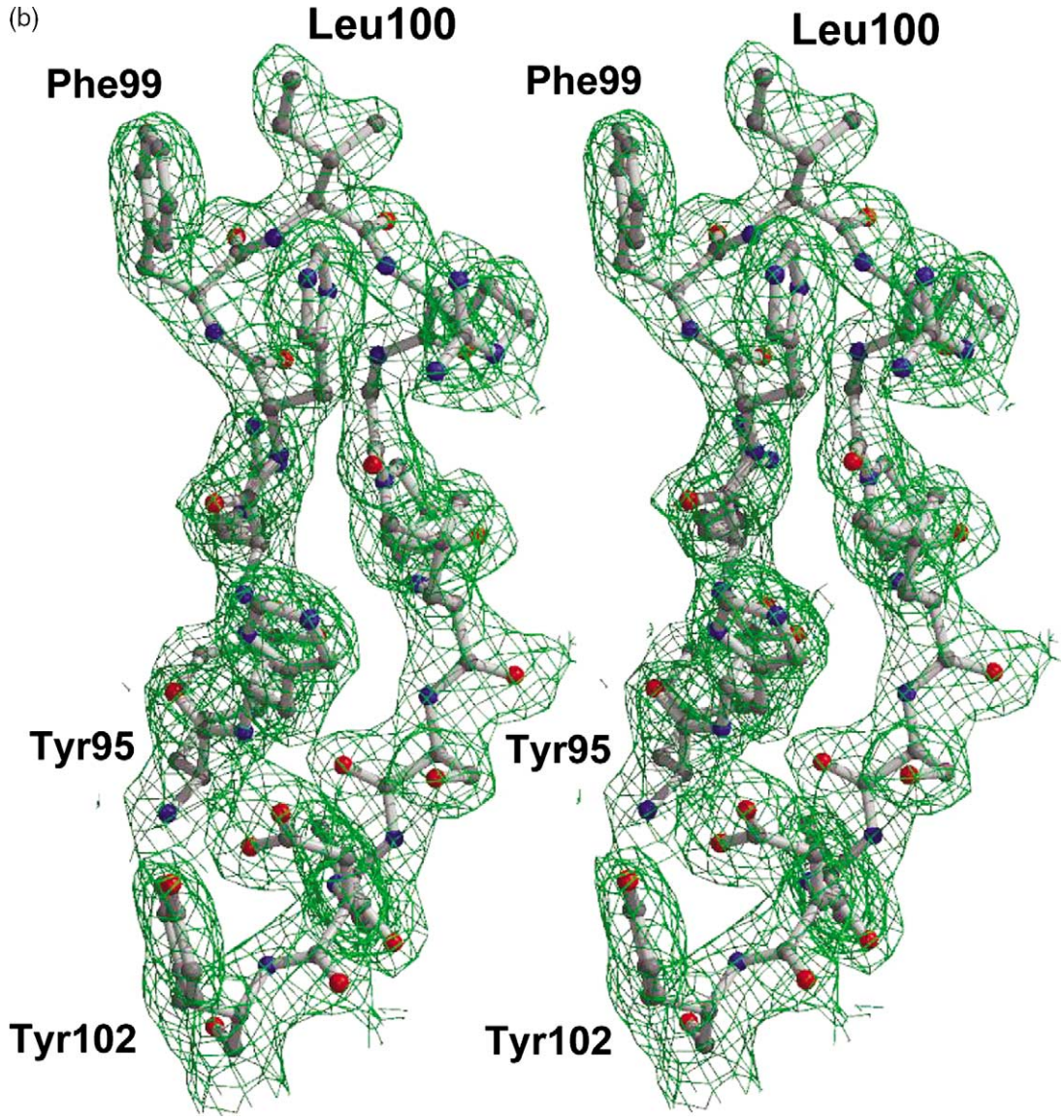

Figure 2. Stereoviews illustrating the electron density maps ($2F_o - F_c$) contoured at 1σ for heavy chain hypervariable loops (a) H2 and (b) H3 of m18. The protein is shown as ball-and-stick models with colored atoms (carbon gray, nitrogen blue and oxygen red) and the electron density as green nets. The Figure was prepared using BOBSCRIPT.[78]

between V_H and V_L, which do not interfere with the binding of lysozyme.[48,49] There are three Fab–Fab interfaces (V_H–C_{H1}) in the crystal structure of m18, and each has one sulfate ion at a conserved location that makes contacts with Asn31 and Tyr33 of H1, and Ser128 and Lys129 of the C_{H1} domain. Similar interactions of sulfate ions were noted at the Fab–ligand[50] and Fab–Fab[51] interfaces.

Canonical structures of CDR L1, L2, L3, and H1

The antigen-binding site of Fab is formed by six hypervariable regions, three from the V_L domain (L1, L2, and L3) and three from the V_H domain (H1, H2, and H3), also called antibody-combining site, which conforms to the epitope. To date, five of the six CDRs have been shown to posses standard main-chain conformations and were classified into different canonical structures.[52] The H3 motif has no standard conformation, due to its variable length. The loops are specified by the hypervariable structural definition and numbered according to the Kabat numbering scheme.[53] In the light chain of m18, L1 (residues L24–L34) is 11 residues long and belongs to canonical structure 2. The canonical structure 2 of the L1 region in the V_κ domain has two forms, A and B, depending upon the peptide conformation between residues L30 and L31. The L1 of m18 has the A form and is packed against the key residue at the framework position 71, which is a conserved phenylalanine residue as found in other V_κ human germline segments.[54] There is a salt-bridge connecting the side-chains of residues L28 (Asp) and L30 (Gln) at the tip of the L1. The side-chains of Ala25, Ile29, and Leu33 point inward within the L1, forming a hydrophobic core that provides stability in lieu of the interstrand hydrogen bonds. L2 (residues L50–L56) belongs to the category of canonical structure 1 and forms a classical γ-turn with the central residue L51 (Ala).

(a)

Heavy chain (V_H)

```
             10        20        30          40          50        60          70
              |         |         |           |           | a       |           |
m18   EVQLLESGPGVVKPSETLSLTCTVSGASVN NYYWT WVRQPPGKGLEWIG NVY-DSGDTNYNPSLSS RLSLS
b12   QVQLVQSGAEVKKPGASVKVSCQASGYRFS NFVIH WVRQAPGQRFEWMG WINPYNGNKEFSAKFQD RVTFT
F105  QVQLQESGPGLVKPSETLSLTCTVSGGSIS SHYWS WIRQSPGKGLQWIG YIY-YSGSTNYSPSLKS RVTIS
      :*** :**. : **. ::.::* .**  .. ..    *:**.**: ::*:* :   .*..:: ...:.. *::::
                                    CDR H1              CDR H2

             80          90        100                  110
              | abc       |          |abcdef              |
m18   MDTSKNQFSLRLSSVTAADTATYYCAR YHRHFIRGPLSF----DY WGRG TLVTVSSA
b12   ADTSANTAYMELRSLRSADTAVYYCAR VGPYSWDDSPQDNYYMDV WGKG TTVIVSSA
F105  VETAKNQFSLKLTSMTAADTAVYYCAR GPVPAVFYGDYRL---DP WGQG TLVTVSSA
       :*: *   :.* *: :****.***** *  **:* * * ****
                                  CDR H3
```

Light chain (V_L)

```
             10        20        30          40          50        60          70
              |         |         |           |           |         |           |
m18   ELQMTQSPSFLSASVGDRVSITC RASQDIQ-KFLA WYQLTPGDAPKLLMY SASTLQS GVPSRFSGSGSGTE
b12   EIVLTQSPGTLSLSPGERATFSC RSSHSIRSRRVA WYQHKPGQAPRLVIH GVSNRAS GISDRFSGSGSGTD
F105  EIVLTQSPGTLSLSAGERATLSC RASQSVSSRYLA WYQQKPGQAPRLLIY GASSRAT GIPDRFSGSGSGTD
      *: :****. ** * *:*.:::* *:*:.:  : :* *** .**:**:*::: ..*.  : *:..*********:
                              CDR L1                CDR L2

             80          90        100
              |          |          |
m18   FTLTISGLQPEDFATYYC QHLKRYPYT FGQGTKLEISR
b12   FTLTITRVEPEDFALYYC QVYGASSYT FGQGTKLERKR
F105  FTLTISRVEPEDFAVYYC QQYDNSVCT FGQGTKLEIKR
      *****: ::****** *** *       * ******** .*
                        CDR L3
```

(b)

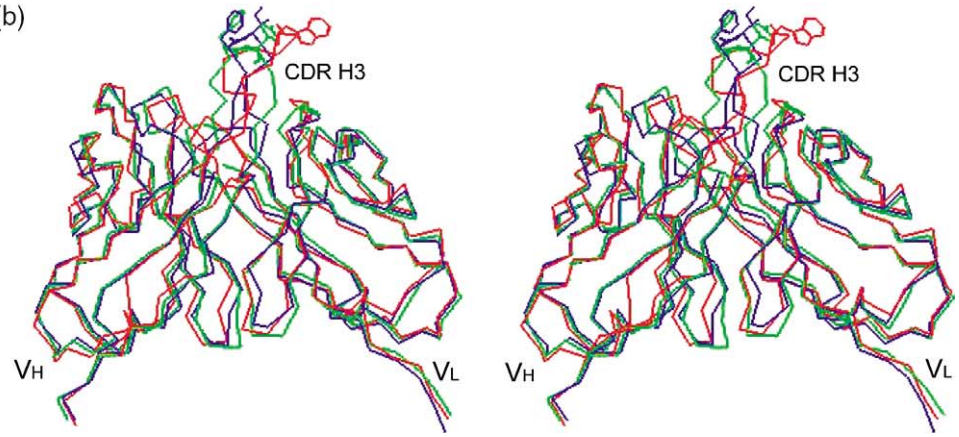

Figure 3. (a) Amino acid sequences of variable domains for the heavy and light chains (V_H and V_L) of Fab m18 are compared with those of the CD4bs antibodies b12 and F105. The six complementarity-determining regions (CDRs) are shown in bold face. (b) The stereographic drawing represents the superposition of the variable domains (V_H and V_L) of m18 (in blue), b12 (in red), and F105 (in green), and shows the aromatic residues on the apex of extended H3s.

Ala51 is flanked by two Ser residues and is found in a strained conformation, having the φ, ψ angles of +70°, −51°. The average φ, ψ values observed for this type of three-residue turn are +75°, −60°.[38] The backbone amide nitrogen atom of residue L51 is engaged in a hydrogen-bonding interaction with the carbonyl oxygen atom of residue L33 from L1. L3 (residues L89–L97) has canonical structure 1 and is most commonly observed in the L3 region of V_κ domains. The L3 motif makes a number of intermolecular contacts with L1, including three consecutive backbone hydrogen bonds between the segments of loops containing residues L89–L91 and L32–L34. It also makes hydrogen bonds involving the side-chains of residues L92 (Lys) and L93 (Arg) to the backbone carbonyl oxygen atoms of residues L27 and L30. In the heavy chain of m18, H1 (residues H31–H35) corresponds to canonical structure 1 and packs across the top of the variable domain. There is a cation–π interaction between Tyr32 and the framework residue Arg94 at the bottom of H3. The backbone amide and carbonyl groups of Trp34 form a hydrogen bond with the carbonyl and amide groups of Val51, respectively,

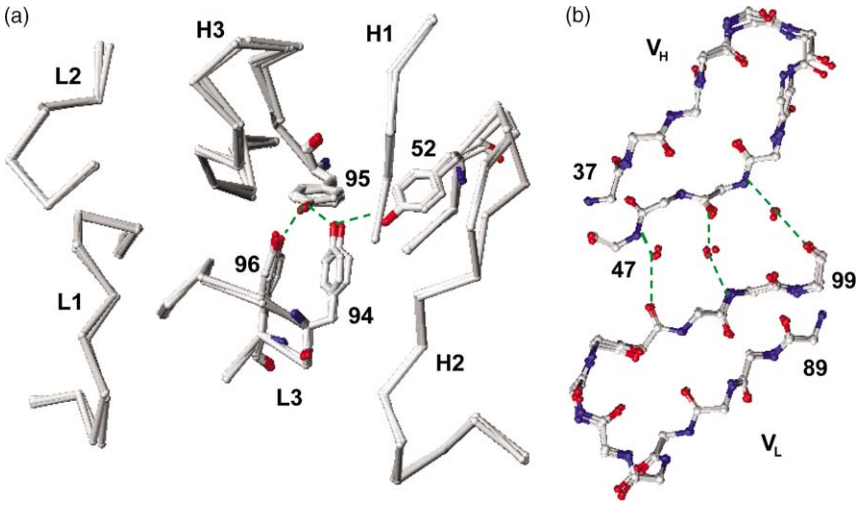

Figure 4. (a) The tyrosine residues at positions H94, H96, H52, and H95 form a quadruple tyrosine motif that connects L3, H2, and H3. (b) Three water molecules buried in the V_H–V_L interface (between two β-hairpins from the V_H/V_L frameworks) are shown as red spheres.

from H2. The interactions of CDR loops between themselves and to the framework regions stabilize the antibody-combining site and might contribute to the specificity of m18–gp120 binding.

Novel conformational features of CDR H2

The H2 motif (residues H50–H65) is 16 residues long. Residues H56–H58 form the short C″ strand, and the variation in the conformation of H2 is limited to residues H52 through H56. Insertions occur at position H52, except for canonical structure 1, and conformations of inserted residues 1–3 lead to different canonical structures of H2. A superimposed view of residues H50–H58, corresponding to the highly variable portion of H2, from the three crystallographically independent Fab m18 molecules is presented in Figure 5(a). As Fab m18 does not have insertion at position 52, it is expected to exhibit canonical structure 1. However, it is significantly different. Although the conformation of H2 appears to be similar to canonical form 1, it has a distinct, bulged β-hairpin structure without any interstrand cross-linking hydrogen bonds, which are present in all four canonical structures of H2. The glycine residue at position 55 is conserved in the three structures used to define canonical form 1.[52] The comparison of H2 structures of similar sizes from different HIV-specific antibodies reveals a markedly different, bulged H2 conformation for m18 (Figure 5(b)). The reason for this bulge may be the presence of Tyr52, which points into the interior of the loop and hence sterically disfavors the formation of interstrand hydrogen bonds (Figure 5(a)). The phenolic oxygen atom of Tyr52 is hydrogen bonded to the amino group of an asparagine residue at position 58.

Intriguingly, we observed that one of the three H2 CDRs in the asymmetric unit exhibits a different conformation in the area of Gly55 (Figure 4(a)). Although this conformational difference results in

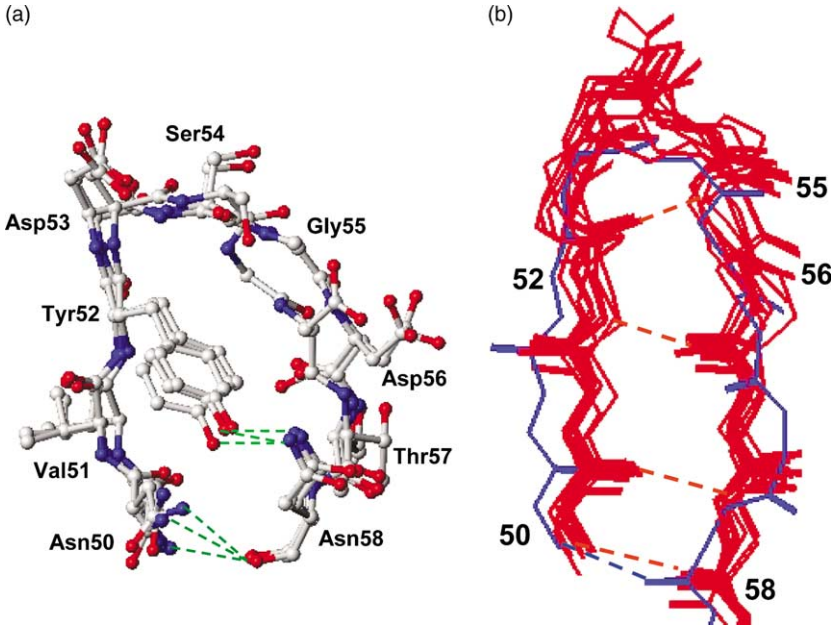

Figure 5. (a) A superimposed view of residues 50–58 of the H2s from the three Fab m18 molecules in the asymmetric unit. Hydrogen bonds are denoted by dotted green lines. (b) Superposition of the H2s from various HIV antibodies based on C^{α} positions. The H2 of m18 is blue and that for other HIV antibodies are red.

a 3.5 Å distance between the C^{α} positions of Gly55 in the two conformations, it does not affect the characteristic hydrogen bond between residues Tyr52 and Asn58, and preserves the overall conformational features of H2 newly observed in the current structure (Figure 5(a)).

To understand the significance of this novel conformation of H2, we analyzed the germline sequences from the IMGT sequence database,[55] and found that the H2 conformation is not due to mutational effects derived from affinity maturation, because residues Try52 and Asn58 are unchanged when compared to the corresponding VH4 gene family. Furthermore, many V_H sequences in the NCBI sequence database[56] share similarities with the H2 sequence of m18. To date, the only structure solved from the related V_H gene is Fab F105,[35] which also contains the tyrosine and asparagine residues at positions 52 and 58, respectively (Figure 3(a)). However, there is no hydrogen bond linking the two residues, and Tyr52 is oriented out of the loop. Therefore, the H2 of F105 belongs to the regular canonical form 1. These findings suggest that the H2 of m18 may form a new sub-class of canonical form 1, whose defining feature is the existence of residue Tyr52, which points into the interior of the H2 loop.

Protruding, rigid, and long CDR H3 with an α-turn motif at the apex

The H3 of m18 (residues H95–H102) is a 14 residue sequence with a six residue insertion after position H100. Unlike many of the antibody structures containing long H3 sequences that usually exhibit severe disorders, the H3 structure of m18 is well defined for all three molecules in the asymmetric unit (for example, see Figure 2(b)). The H3 motif, containing predominantly hydrophobic residues, forms nearly a β-sheet structure with a five residue α-turn at the apex where a β-turn is most frequently observed. All side-chains of H3 residues in the three m18 molecules in the asymmetric unit have similar conformations (Figure 6). The large variations in sequence, size, and structure of H3s present difficulties in defining canonical structures.[57] However, certain empirical rules for predicting the conformation of H3 on the basis of the nature and location of key amino acid residues have been formulated.[58–61] The key residues of framework at positions H93, H94, and H101 are occupied by Ala, Arg, and Asp, respectively, which predict a kinked base for H3 according to the H3 rules. This prediction is consistent with the H3 structure of m18, in that the dihedral angle formed by four consecutive C^{α} atoms ending with H103 at the base is acute; the average value of this angle is 25° for the three molecules of m18, corresponding to the kinked (0–60°) structure. The same set of key residues appears in the b12 and F105 structures, forming dihedral angles of 12.4° and 24.7°, respectively, which also indicate kinked bases. The structural features of H3s observed in the antibodies m18, b12, and F105 are compared in Figure 7. Apart from the kinked bases, the tips of the three H3s have prominent hydrophobic residues. Although the H3 of m18 is shorter than that of b12 by four residues, the height and extent of projection from the antibody-combining sites are very similar; the distances between the C^{α} atoms of H94 at the base and H100 at the tip are 17.5 Å and 18.2 Å for m18 and b12, respectively. Notably, this distance is only 13.3 Å for the H3 of F105, although it is one residue longer than that of m18. This is due to the fact that bulges are present in the H3 of b12 and F105, whereas the H3 of m18 has an extended structure, despite the presence of a proline residue at H100C. Notably, the H3 of F105 has three proline residues that might influence its conformation. Among the three antibodies, the H3 of m18 has cross-linking interstrand hydrogen bonds as observed in a regular β-sheet structure, which might exclusively confer rigidity to the loop. We observe also that the side-chains of hydrophobic (Phe, Leu, and Pro) and basic (His and Arg) residues emanating from alternating sides of the sheet pack against each other and act like a zipper, which may provide more rigidity in addition to the

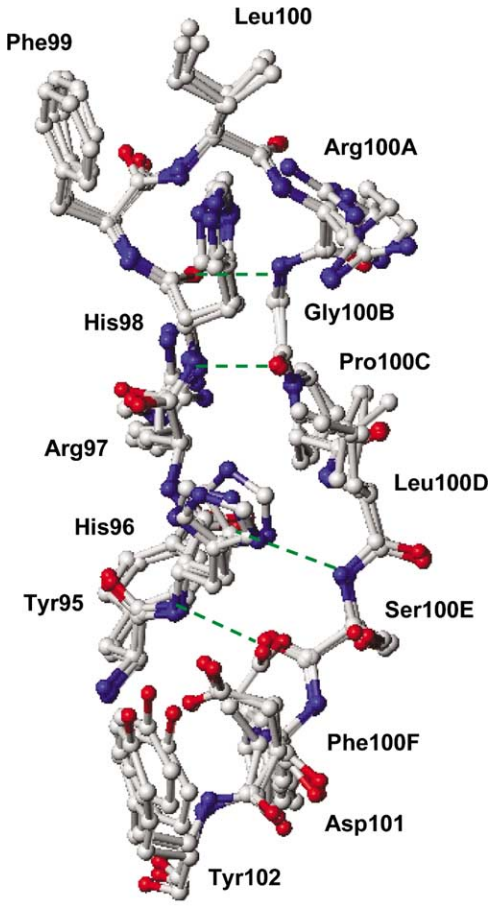

Figure 6. Superposition of m18 H3s of the three Fab molecules in the asymmetric unit. The dotted lines represent hydrogen bonds.

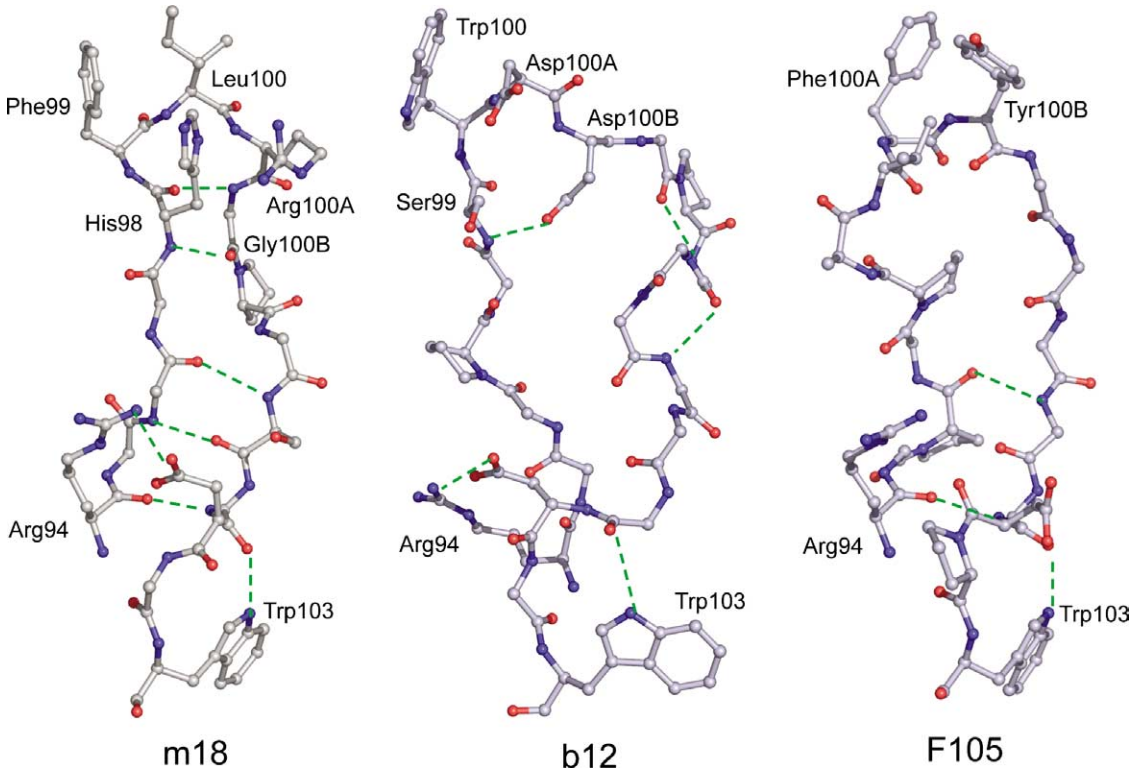

Figure 7. Comparison of the conformation and hydrogen bonding pattern of the H3s from m18, b12, and F105. Residues Arg94 and Trp103 from the framework regions play key roles in maintaining the H3 conformations according to the H3 rules. Amino acid residues at the apex of the loops are labeled and the side-chains of other residues are omitted for clarity.

interstrand hydrogen bonds (Figure 6). All these interactions stabilize the H3 conformation, providing a high degree of rigidity. No significant lattice contact is observed. Together, H3 appears to have a spearhead at the crown of the loop bearing a phenylalanine-leucine fork. This protruding and rigid H3 loop of m18 antibody might have the potential to plunge into the recessed CD4bs on gp120 (details of the proposed recognition mechanism using docking models of gp120-m18 complexes are described later).

Structural mimicry of the CDR2-like C′C″ region of CD4 by the H3 of Fab m18

The most remarkable feature of the m18 structure is the dominance of H3 in the antibody-combining site, which adopts nearly a β-hairpin conformation and closely resembles the Ig CDR2-like C′C″ region of CD4 first domain D1.[62] The C′C″ region of CD4 spanning residues 40–48 accounts for 63% of interatomic contacts with gp120, where Phe43 at the tip of the loop itself contributes up to 23%.[26] The H3 of m18 also contains a phenylalanine residue (Phe99) exposed at the tip, which closely mimics the hotspot residue Phe43 of the C′C″ loop in CD4 (Figure 8). The H3 of m18 has a five residue α-turn motif, which is sturdy and invariably has a functional role in molecular recognition.[63] The dominant structural role of Phe43 of CD4 as a hotspot in the binding of gp120 suggests a likely functional role for m18 residue Phe99 as a CD4 mimetic. The striking similarities at the phenylalanine residue and robust β-sheet features observed in these loops suggest possible protein grafting of H3, mimicking the CDR2-like C′C″ loop

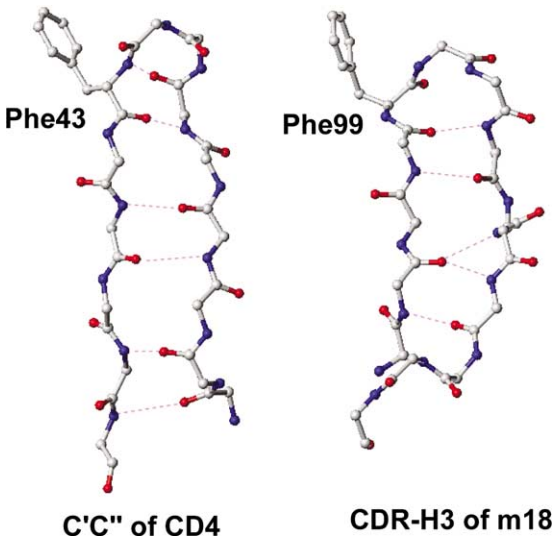

Figure 8. Backbone skeletal representations of the C′C″ loop of CD4 and the H3 of m18 indicate β-hairpin motifs as a common template, with a phenylalanine residue exposed at the tip.

of CD4, which might offer a strategy for developing antibody-based CD4 mimetics to inhibit HIV entry.

Docking of Fab m18 to CD4-bound and CD4-free gp120

To assess the ability of m18 H3 to approach the recessed CD4bs and to understand the structural mechanisms of m18 binding to gp120, we performed docking studies based on the gp120–CD4 interfacial cavity information[26] and the locations of conserved neutralizing epitopes overlapping the CD4bs of gp120.[11,18] The knowledge of conformational changes at CD4bs from the known crystal structures of the HIV-1–gp120 complex[26] and the unliganded simian immunodeficiency virus (SIV) gp120,[36] which have more than 70% sequence similarity to HIV-1, along with the established rules for antigen–antibody interactions that are limited to the CDRs of the antibody,[64] offers the opportunity to explore the probable binding modes in the m18-gp120 complex. Using the *FTDOCK* step in the 3D-Dock program, 10,000 docked complexes were obtained on the basis of the surface and charge complementarity scores. The *rpscore* routine, which scored and sorted the docking solutions, was used to calculate the residue pair potentials. The final step of 3D-Dock is *filtering*, for which we used an intermolecular distance constraint on the basis of available biochemical information. In the bound gp120 crystal structure, CD4bs is located in a depression formed at the interface of outer and inner domains with the bridging domain. From earlier mutagenesis studies, it was found that the changes in Asp368 and Glu370 uniformly disturbed the binding of CD4bs antibodies to gp120, and the change in Trp427 affected the binding of only a few CD4bs antibodies, including b12.[11,18] We measured the binding of gp120 alanine-scan mutants to m18 and correlated the conservation of amino acid residues in 380 isolates with the relative energy of m18 binding (Table 2). One might expect that the m18 H3 plays a role in gp120 recognition. Thus, a single intermolecular distance constraint involving residue Trp427 at the bottom of recessed CD4bs and residue Phe99 at the tip of H3 was employed and possible docked orientations for the m18–gp120 complexes with bound and free forms of gp120 were identified (depicted in Figure 9(a) and (b), respectively). As shown in Figure 9(a), Fab m18 binds to the bound form of gp120 in approximately the same region as CD4; i.e. between the outer and inner domains, which are connected by a four-stranded bridging sheet. The Fv portion of m18, containing two immunoglobulin domains V_H and V_L, rotates $\sim 45°$ with respect to the CD4 orientation, thereby increasing the geometric fit and avoiding steric clashes due to its increased size compared to the D1 domain of CD4. A total solvent-accessible surface area of 2080 Å^2 is buried in the m18–gp120 interface (1010 Å^2 from gp120 and 1070 Å^2 from m18) and these values are similar to those in the docking model of the gp120-b12 complex (total surface 2070 Å^2, including 1030 Å^2 on gp120 and 1040 Å^2 on b12).[22] Figure 9(b) illustrates the mode of m18 binding to CD4-free gp120 as seen in the recent crystal structure of SIV gp120.[36] As shown in Figure 9(a) and (b), the m18 H3 can reach the center of the outer domain of both free and bound forms of gp120 without steric restraint. In the docking model of the free gp120–m18 complex, a total interface area of 1460 Å^2 is buried between gp120 and m18 (680 Å^2 from gp120 and 780 Å^2 from m18), which is similar to the 1540 Å^2 of the gp120–CD4 interface (800 Å^2 from gp120 and 740 Å^2 from CD4).[26] On the basis of these docking models, we suggest that the CD4-binding loop in the bound as well as the free forms of gp120 could contact the H3 of m18.

Comparisons of the binding sites of m18 and CD4 on gp120

The Fab m18 binding sites on gp120 in the two docked complexes are superimposed in Figure 10(a). Contact residues (distance cutoff $\leq$4.5 Å) at the binding interface were identified (the complete list of contact residues in the two docked complexes is given in Table 3). A common interaction pattern is found in the two m18-gp120 docking models (Figure 10(a)), which is observed also in the gp120–CD4 structure (Figure 10(b)). Residue Phe99 at the tip of H3 approaches the backbone atoms of Trp427 in gp120 (Trp440 in unliganded SIV gp120, Figure 10(a)) in a manner similar to the CD4 hotspot residue Phe43 (Figure 10(b)). Furthermore, the m18 H3 residue Arg100A makes a hydrogen bond with Asp368 of gp120 (Asp384 for unliganded SIV gp120) (Figure 10(a)), which mimics another hotspot interaction between Arg59 of CD4 and Asp368 of gp120 (Figure 10(b)). Residues Glu386 and Val430 are located at the interface of the docked m18–gp120 complexes (Figure 10(a)) and are present also at the CD4–gp120 interface (Figure 10(b)). The solvent-accessible area surrounding the Phe43 cavity in the docked m18–gp120 complex (with the gp120 bound form) is shown in Figure 10(c) and compared with the CD4-gp120 structure (Figure 10(d)). In agreement with the docking model, alanine mutation of Asp368, Pro369, and Glu370 in HIV-1 gp120 completely abolished binding to m18 (Table 2) as well as to CD4bs antibodies b12 (except for Pro369) and m14.[11,20] The residues identified from mutagenesis studies (Gly367, Asp368, Pro369, Glu370, Lys429, Asp474, and Met475 in HIV-1 gp120 and the equivalent SIV gp120 residues Asp384, Glu386, and Lys442) along with Pro385 appear at the interface. Many of them are highly conserved (Table 2) and make direct protein–protein contacts in the docked m18–gp120 complexes (Table 3). The same set of gp120 residues in direct contact with m18 has been identified in the two docked complexes, including Asp368, Glu370, Trp427, Lys429, and Val430 in HIV-1 and Asp384, Glu386, Trp440, Lys442, and Val443 in SIV (Table 3). Among these residues, Trp427 and

Table 2. Binding of m18 to alanine scan mutants of JR-CSF gp120

Mutant[a]	gp120 domain[b]	Conservation (%)[c]	Relative affinity[d]	Mutant[a]	gp120 domain[b]	Conservation (%)[c]	Relative affinity[d]
Wild-type			100	P417A	C4 (V4 base)	79	92
C119A		99	48	R419A	C4 (V4 base)	81	56
V120A		98	56	I420A	C4	97	150
K121A	C1(V1/V2 stem)	91	114	**K421A**	C4	91	38
L122A	C1(V1/V2 stem)	94	72	Q422A	C4	98	112
T123A[e]	C1(V1/V2 stem)	99	59	I423A	C4	92	251
L125A[e]	C1(V1/V2 stem)	98	170	I424A	C4	65	163
V127A	C1(V1/V2 stem)	99	135	N425A[e]	C4	85	109
T198A	C2(V1/V2 stem)	86	70	M426A[e]	C4	82	118
S199A[e]	C2(V1/V2 stem)	94	32	W427A[e]	C4	98	104
V200A	C2(V1/V2 stem)	42	136	Q428A[e]	C4	95	118
I201A	C2(V1/V2 stem)	89	88	**E429A**[e]	C4	40	14
T202A	C2(V1/V2 stem)	76	85	V430A[e]	C4	86	274
Q203A	C2	99	66	G431A	C4	96	151
A204G	C2	97	237	K432A	C4	42	88
K207A[f]	C2	98	25	M434A	C4	84	95
S256A[e]	C2	97	75	Y435A[e]	C4	99	164
T257A[e]	C2	99	48	P437A	C4	94	96
R298A	C2	99	129	**R469A**	V5	97	35
W338A	C3	98	678	**P470A**	V5	98	38
N339A	C3	72	49	G471A	V5	84	136
P363A	C3	31	4	**G472A**	C5	98	25
S365A[e]	C3	85	69	G473A[e]	C5	98	119
G366A[e]	C3	98	62	**D474A**[e]	C5	71	18
G367A[e]	C3	99	45	**M475A**	C5	90	26
D368A[e]	C3	99	0	R476A	C5	74	258
P369A[e]	C3	42	0	D477A	C5	94	59
E370A[e]	C3	99	0	**W479A**	C5	99	22
Y384A	C3	98	3	ΔV1	Δ134–154		212
N386A	C3	90	3	ΔV1/V2	Δ134–154/ Δ160–193		67
N392A	V4	93	33	ΔV3	Δ303–324		394

[a] Residue numbering scheme is based on the sequence of prototypic HxBc2 gp120 glycoprotein. Mutants with more than a twofold decrease of m18 are highlighted in bold face.
[b] C, constant domain; V, variable loop.
[c] Conservation was calculated as a percentage of the HIV-1 isolates that had the same residue at the same position with respect to a total of 380 isolates sequenced.
[d] Calculated using the formula [apparent affinity (wild type)/apparent affinity (mutant)]×100%, where apparent affinities were calculated as the antibody concentration at 50% of maximal binding.
[e] Residues that exhibit decreased solvent accessibility in the presence of sCD4 (D1D2) in the ternary complex.
[f] Residues involved in maintaining the overall structure of gp120.

Val430 are involved also in the binding of CD4. The gp120 residues shown in italics in Table 3 denote the critical residues for m18 binding, as indicated by the mutagenesis experiment, and those underlined participate in CD4 binding according to the gp120-CD4 structure. These observations indicate that the epitope of Fab m18 significantly overlaps the CD4bs on gp120. The main difference between the two docked complexes is the involvement of the V1-V2 stem region for the binding of m18. In the unliganded form, the V1-V2 stem is recessed in the cavity and does not make any contact with m18. However, for the bound form of gp120, residues from the V1-V2 stem are involved in the binding of m18. This is likely the correct mode of interaction, because one of the residues in the V1-V2 stem is implicated also in the CD4 binding, and antibody b12 is also sensitive to the V1-V2 mutations; binding of CD4 or other CD4 mimics moves the tip of V1-V2 stem over a distance of approximately 40 Å.[26,65]

According to the results from critical assessment of predicted interactions (CAPRI) experiments, most of the failures in docking predictions are mainly due to unexpected, large conformational changes that occur during the binding process.[66] In the case of gp120, it has been shown that conformational change upon the binding of CD4 has a footprint similar to that of CD4bs

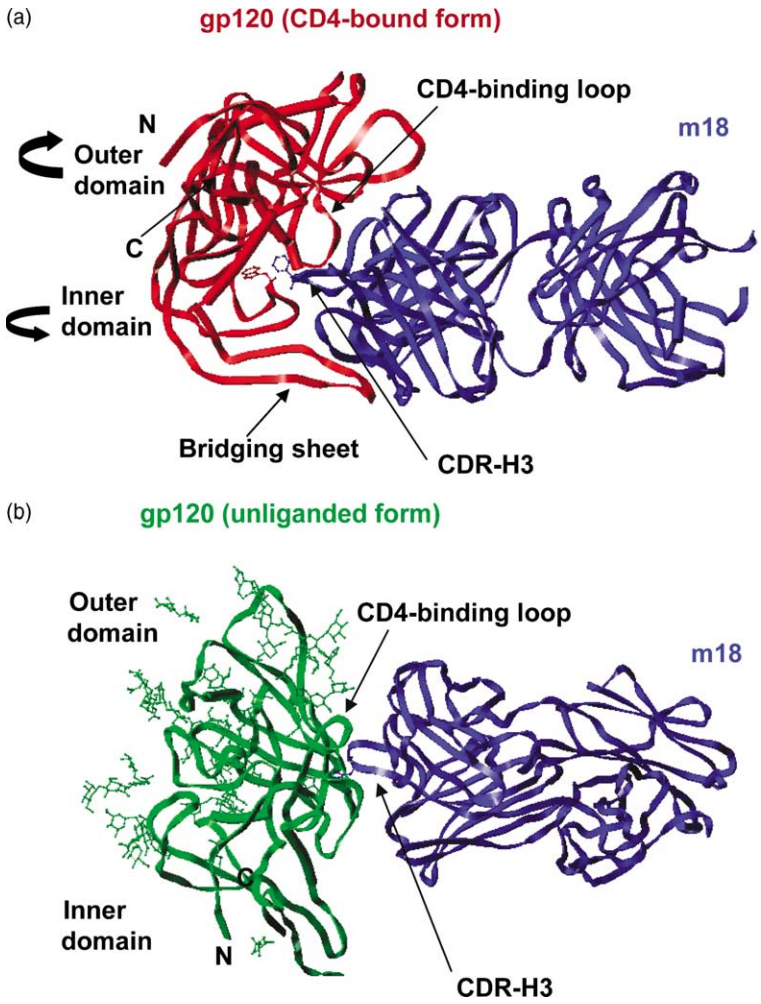

Figure 9. (a) Docking of m18 onto the CD4-bound form gp120. The gp120 molecule is shown in red and m18 in blue. Residues Trp427 of gp120 and Phe99 of m18 H3 are shown as ball-and-stick models. (b) Docking of m18 onto unliganded gp120. The gp120 molecule is shown in green and m18 in blue. Residues Phe99 of m18 H3 and Trp440 of gp120 are shown as ball-and-stick models.

antibodies. Biochemical data generally address the binding residues for one component and often lack the pair-wise contacts or mode of binding between the two. But in the antibody–antigen complex, antibody binds to antigen mainly through the hypervariable loops, which gives reasonable constraints in the docking experiments. Applying such constraints by specifying the CDR residues in the intermolecular distance filter was found successful in the docking predictions for 13 out of 15 antibody–antigen complexes.[67] Rarely, antibodies employ lateral contacts with mostly framework regions, as found in the two camel antibodies with the V_{HH} domain in complex with α-amylase, for which the antibody–antigen interactions could not be predicted correctly.[66] For m18, the H3 motif is long and protruding to reach deep into the cleft formed at the CD4bs. As is shown in Table 3, m18 could use different residues in the CDR loops, specifically in H3 that mimics the CDR2-like C′C″ loop of CD4. Most of the paratope residues of m18 are predicted to be exposed and involved in the antigen binding from contact analysis.[64]

The proposed m18–gp120 model complex has a good structural complementarity and is in agreement with mutagenesis data that identify neutralizing epitopes on gp120. The model, along with biochemical data, implies that Fab m18 recognizes a conserved neutralizing epitope at the gp120 surface which, in part, overlaps the CD4bs of HIV-1 gp120. The availability of gp120 crystal structures in free and bound forms, and in complex with CD4 mimics, rationalizes the requirement for appropriate conformational changes in gp120 upon m18 binding. Antibodies often do not undergo significant conformational changes upon antigen binding, except for changes in the side-chains and flexible loops. Although the conformational changes and binding modes need to be verified by further analysis and experiments, the present docking model is compatible with biochemical data and useful in explaining the recognition mechanisms. Further, the model shows the potential of m18 H3 to approach the recessed CD4bs of gp120, and suggests strategies for optimizing binding affinity and developing antibody-based CD4 mimetics for gp120 binding.

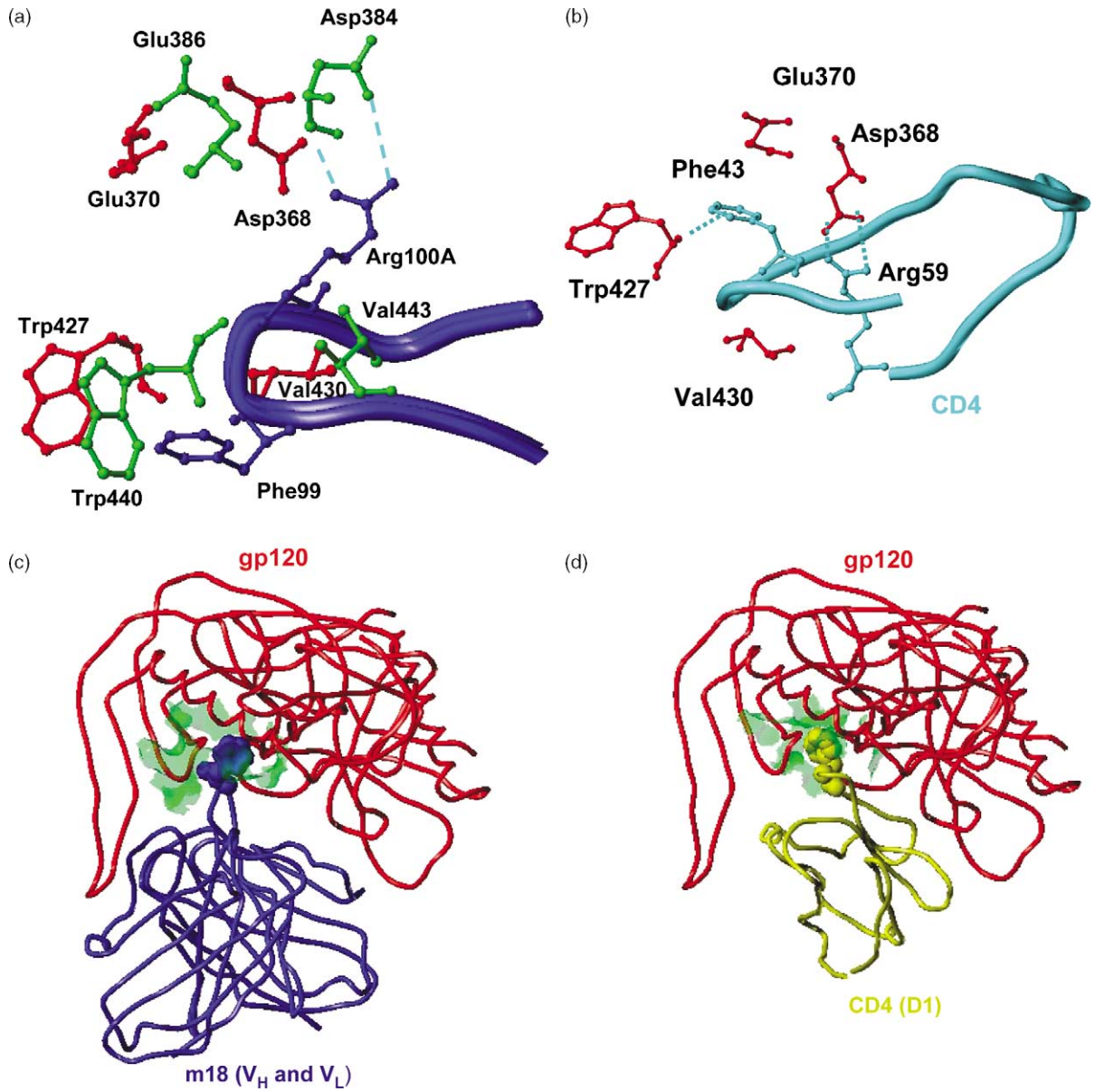

Figure 10. Comparisons of the binding sites of m18 and CD4 on gp120. (a) Superposition of the m18 binding sites from the two docked m18-gp120 model complexes. The binding residues of bound gp120 and free gp120 in the docking models are shown in red and green, respectively; the m18 H3 and the side-chains of Phe99 and Arg100A in the docking models are in blue. (b) The CD4bs on gp120 as observed in the complex structure is depicted. The C′C″ loop of CD4 is shown in cyan and gp120-binding residues in red. The dotted lines represent the intermolecular interactions between the residues. (c) and (d) The Phe43 cavity in m18 (in blue) and that in the CD4 D1 domain (in yellow) are surrounded by the accessible surfaces on the gp120 as observed in the m18-gp120 docking model and the CD4-gp120 crystal structure, respectively. Residues Phe99 of m18 and Phe43 of CD4 are shown as space-filling models.

Conclusions

We have determined the crystal structure of Fab m18, an HIV-specific CD4bs antibody, at high resolution, and found novel conformational features in the H2 and H3 motifs. The H2 motif has a bulged conformation without interstrand hydrogen bonds and represents a sub-class of a canonical structure. The rigid and protruding H3 adopts a β-hairpin-like structure with a phenylalanine residue at the apex. A quadruple tyrosine motif, conserved water molecules, and sulfate ions have been identified at the domain interfaces of the m18 structure. The m18 H3 is strikingly similar to the Ig CDR2-like region of the CD4 D1 domain and contains residue Phe99 mimicking the hotspot residue Phe43 of CD4, which plays a critical role in the formation of the gp120–CD4 complex. Docking simulations of the m18-gp120 complex, taking into account experimental mutagenesis data, predict significant resemblance of the interactions observed in the gp120–CD4 complex. These results suggest that m18 mimics some structural features of CD4, and predict a decreased likelihood for gp120

Table 3. Contact residues within 4.5 Å at the interfaces between m18 and gp120 in the docked complexes of m18 with CD4-bound and CD4-free gp120, respectively

Bound gp120 (28,1005 Å^2)	m18 (25,1074 Å^2)
His105, Ile109 (α1) Cys126, Val127, Gly128, Ala129, Gly194, Ser195, Cys196 (V1V2) Asp279, Asn280, Ala281 (LD) Gly366, *Gly367*, **Asp368**, **Glu370** (β15) Met426, **Trp427**, Gln428, **Lys429**, **Val430** (β20-β21) Gly459, Asn460, Asn462 (V5) Gly473, *Asp474*, *Met475*, Arg476 (α5)	**Lys31** (CDR-L1) Tyr49 (V_H-FR2) Lue54, Gln55, Ser56, Gly57, Val58 (CDR-L2) Pro59, Ser60, Arg61 (V_H-FR3) Arg93, Tyr94, Pro95 (CDR-L3) Thr57, Asn58, Tyr59, Asn60, Pro61 (CDR-H2) **Arg97**, **His98**, **Phe99**, **Ile100**, **Arg100A**, Gly100B, Pro100C (CDR-H3)
Free gp120 (17,681 Å^2)	**m18 (16,789 Å^2)**
Asp384, *Pro385*, **Glu 386**, Val387, Thr388, Phe389 (β15- α3) Tyr400 (β17) Arg434, Ile436, Asn438, Thr439, **Trp440**, His441, **Lys442**, **Val443**, Lys445, Val447 (β20-β21)	**Lys31** (CDR-L1), Ser50, Ser52, Thr53 (CDR-L2) Asn30, Asn31, Tyr32, Tyr33 (CDR-H1) Asp53 (CDR-H2) His96, **Arg97**, **His98**, **Phe99**, **Ile100**, **Arg100A**, Leu100D (CDR-H3)

Number of contact residues and the buried surface area (Å^2) from each interacting partner upon complex formation are given in parentheses. The binding residues that are the same in the two docking models of gp120–m18 complexes are shown in bold face. The gp120 residues identified as critical for binding to m18 from mutagenesis studies (Table 2) are given in italics and those that are involved in the CD4 binding as observed in the crystal structure of gp120-CD4-17b complex are underlined.

escape mutants compared to other antibodies with same affinity but recognizing epitopes that are dissimilar to the CD4bs. Such escape mutations, which affect the binding of m18 and CD4 simultaneously, could lead to lower binding affinity of gp120 to both m18 and CD4, and the virus could lose the ability to replicate if the reproduction ratio falls below 1. In an ideal situation of complete mimicry, the energy profile of gp120–antibody binding would be identical with that of the receptor, and mutations that lead to neutralization escape mutants would adversely affect the binding to CD4 and the efficacy of virus entry. Thus, m18 and similar antibodies could have potential as effective HIV inhibitors.

Materials and Methods

Preparation of anti-HIV Fab m18, crystallization and data collection

The Fab m18 was previously selected from a human phage-display library by sequential antigen panning (SAP) against different soluble HIV-1 Envs and the Env complexes with soluble CD4 (sCD4).[21] The Fab was found to bind with high affinity to different Envs, and exhibited cross-reactive HIV neutralizing activity. The anti-HIV Fab m18 was produced by standard procedures[68] and a protein G (Amersham) column was used for purification. The m18 protein was finally exchanged in 50 mM Tris–HCl (pH 7.5) with 100 mM NaCl and concentrated to 10 mg/ml. The alanine-scan mutants of gp120 and the assay used to measure their binding to m18 have been described,[11] and the experiments were performed in Dennis Burton's laboratory (The Scripps Research Institute).

Initial screening of crystallization conditions was performed with several screens from Hampton Research (Laguna Niquel, CA) and Wizard Screen (deCODE genetics, Bainbridge Island, WA). A 'Hydra II Plus' (Matrix Technologies, Hudson, NH) crystallization robot was employed to set the screens with the sitting-drop, vapor-diffusion method at room temperature. Rod-shaped crystals of m18 from 100 mM Ches buffer (pH 9.5) and 20% (w/v) PEG-8000 appeared in two to three days. The diffraction data were collected at the home laboratory using a Rigaku rotating anode X-ray generator and a MAR345 image plate. The crystal was found suitable for X-ray diffraction with cryoprotectant consisting of well solution and 20% glycerol and diffracted to 2.03 Å resolution. Data were processed and scaled with the HKL2000 program.[69] The crystal belongs to space group $P2_1$ with unit cell dimensions $a=48.59$ Å, $b=82.62$ Å, $c=187.20$ Å, and $\beta=95.26°$. There are three Fab molecules in the asymmetric unit and the Matthews coefficient is estimated to be 2.6 Å^3 Da^{-1}, corresponding to a solvent content of 52.4% (v/v).

Structure solution and refinement

The structure of Fab m18 was solved by molecular replacement using AMoRe.[70] The amino acid sequences of constant domains (C_L and C_{H1}) of m18 are identical with that of b12[22] and X5,[28,29] except for two residues in each domain. But, the variable domain of m18 shares only about 50% of sequence identity with b12 and X5. The initial attempts at finding a solution using the whole Fab molecule of either b12 or X5 as search models failed. Since the variable and constant domains in each chain of Fab are connected by a flexible elbow region, the relative orientation of the two domains can vary. In such cases, the molecular replacement attempts with search models of individual domains have been proven successful.[71] For Fab m18, the constant domains of both light and heavy chains (C_L and C_{H1}) of b12 were used as the search model. Three distinct solutions corresponding to the three Fab molecules in the asymmetric unit were obtained, two of them had the same correlation coefficient (22.9) and R value (50.9), while the correlation coefficient and R value for the third solution were 22.3 and 51.0, respectively, for the X-ray data ranging from 15 Å to 2.5 Å. Crystal packing analysis showed the coherent ordering of the three molecules containing constant domains and possible spatial locations for the variable domains without any clashes.

The initial model with three molecules of constant domains was subjected to rigid body refinement (30–2.03 Å) and the *R* value was 45.7% at that stage. As mentioned earlier, the variable domains of m18 share a sequence similarity to b12 of 46.3%. Interestingly, a BLAST search against m18 F_v sequence reveals that the F_v portion of a human IgM cold agglutinin, of which the structure is available, shares 64.2% identical sequence with the m18 F_v.[72] The calculated electron density maps were of good quality and facilitated the construction of variable domains with the guidance of the cold agglutinin Fv structure. Model building was performed with O,[73] and subsequent refinement with CNS.[74] Each cycle of refinement was done with global *B*-value corrections and bulk solvent corrections. Several cycles of refinement and model building allowed proper fitting of the side chains of F_v residues and *de novo* modeling of CDR loops on the basis of $2F_o-F_c$ and F_o-F_c electron density maps. When the refinement was about to complete with an *R* value of 25%, water molecules were included by the water-picking routine of CNS,[74] and reasoned with electron density and geometric criteria. Three sulfate anions from the Ches buffer were identified unambiguously in the Fab–Fab interfaces from the F_o-F_c map. Cross-validation was carried out with a randomly selected test data set of 4.3% of the total number of reflections. The programs PROCHECK[75] and WHATIF[76] were used to assess the stereochemical quality of the final model.

Docking simulations of the gp120–m18 complexes

The Fab m18 was docked onto the CD4bs of gp120 in both CD4-bound and CD4-free forms using the 3D-Dock suite of programs.[77] The conformation of unliganded SIV gp120 was assumed to be the conformation of HIV gp120 in its CD4-free form for the purpose of docking predictions. The gp120 molecule was considered as the static unit and m18 as the mobile unit. The docking procedure was composed of three steps. First, the FTDOCK module globally scanned rotational and translational space (with the use of default grid size 234, surface thickness 1.3 Å, and angle step 3°) for possible orientations of the two molecules (gp120 and m18) in the complex, which were limited by surface complementarity and electrostatic scores, resulting in a total of 10,000 different complexes. Second, empirical scores for the model complexes, using residue level pair potentials (rpscore), were calculated. Third, the filtering of complexes was performed on the basis of the structural and biochemical data. In both docking simulations, the intermolecular distances involving the m18 epitope on gp120 and the H3 residues of m18 were incorporated into the filtering step as constraints. Specifically, the distance constraint between residue Trp427 of the bound form gp120 (Trp440 in the free form SIV gp120) and residue Phe99 at the tip of the m18 H3 was used as an intermolecular constraint for the identification of a gp120–m18 complex. In the docking experiment of m18 with the bound form of gp120, the distance filter involving the H3 residue Phe99 of m18 and the entire gp120 structure brought the number of predictions from 10,000 to 1073. Further filtering with the involvement of Trp427 selected only two complexes, one of which had a positive rpscore value of 0.870 and was ranked the 160th among the 10,000. The same filtering procedure was applied in the docking experiment of m18 with the free form of gp120, resulting in seven complexes, of which the top solution was ranked the 42nd among the 10,000. The conformation of side-chains at the interface was optimized for favorable contacts with the use of a rotamer library in the program O.[73]

Acknowledgements

We thank D. Burton for providing the gp120 mutants and for help with the alanine scan mutagenesis. We thank I. Sidorov for his efforts in antibody production, S. Ravichandran for help in using the docking program, and P. Johnson for critical reading of the manuscript. We thank the Advanced Biomedical Computing Center of NCI-Frederick for computing facilities and the IATAP of NIH for support to D.S.D. This research was supported by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, and by Federal funds from the National Cancer Institute, National Institutes of Health, under contract no. NO1-CO-24000 and NO1-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does the mention of trade names, commercial products, or organizations imply endorsement by the US Government.

References

1. Dimitrov, D. S. (2004). Virus entry: molecular mechanisms and biomedical applications. *Nature Rev. Microbiol.* **2**, 109–122.
2. Smith, A. E. & Helenius, A. (2004). How viruses enter animal cells. *Science*, **304**, 237–242.
3. Burton, D. R. (2002). Antibodies, viruses and vaccines. *Nature Rev. Immunol.* **2**, 706–713.
4. Ferrantelli, F. & Ruprecht, R. M. (2002). Neutralizing antibodies against HIV—back in the major leagues? *Curr. Opin. Immunol.* **14**, 495–502.
5. Veazey, R. S., Shattock, R. J., Pope, M., Kirijan, J. C., Jones, J., Hu, Q. *et al*. (2003). Prevention of virus transmission to macaque monkeys by a vaginally applied monoclonal antibody to HIV-1 gp120. *Nature Med.* **9**, 343–346.
6. Barbas, C. F., III, Collet, T. A., Amberg, W., Roben, P., Binley, J. M., Hoekstra, D. *et al*. (1993). Molecular profile of an antibody response to HIV-1 as probed by combinatorial libraries. *J. Mol. Biol.* **230**, 812–823.
7. Burton, D. R., Barbas, C. F., III, Persson, M. A., Koenig, S., Chanock, R. M. & Lerner, R. A. (1991). A large array of human monoclonal antibodies to type 1 human immunodeficiency virus from combinatorial libraries of asymptomatic seropositive individuals. *Proc. Natl Acad. Sci. USA*, **88**, 10134–10137.
8. Burton, D. R., Pyati, J., Koduri, R., Sharp, S. J., Thornton, G. B., Parren, P. W. *et al*. (1994). Efficient neutralization of primary isolates of HIV-1 by a recombinant human monoclonal antibody. *Science*, **266**, 1024–1027.
9. Roben, P., Moore, J. P., Thali, M., Sodroski, J., Barbas, C. F., III & Burton, D. R. (1994). Recognition properties of a panel of human recombinant Fab fragments to the

CD4 binding site of gp120 that show differing abilities to neutralize human immunodeficiency virus type 1. *J. Virol.* **68**, 4821–4828.
10. Zwick, M. B., Parren, P. W., Saphire, E. O., Church, S., Wang, M., Scott, J. K. *et al*. (2003). Molecular features of the broadly neutralizing immunoglobulin G1 b12 required for recognition of human immunodeficiency virus type 1 gp120. *J. Virol.* **77**, 5863–5876.
11. Pantophlet, R., Ollmann, S. E., Poignard, P., Parren, P. W., Wilson, I. A. & Burton, D. R. (2003). Fine mapping of the interaction of neutralizing and nonneutralizing monoclonal antibodies with the CD4 binding site of human immunodeficiency virus type 1 gp120. *J. Virol.* **77**, 642–658.
12. Parren, P. W., Fisicaro, P., Labrijn, A. F., Binley, J. M., Yang, W. P., Ditzel, H. J. *et al*. (1996). In vitro antigen challenge of human antibody libraries for vaccine evaluation: the human immunodeficiency virus type 1 envelope. *J. Virol.* **70**, 9046–9050.
13. Ho, D. D., McKeating, J. A., Li, X. L., Moudgil, T., Daar, E. S., Sun, N. C. & Robinson, J. E. (1991). Conformational epitope on gp120 important in CD4 binding and human immunodeficiency virus type 1 neutralization identified by a human monoclonal antibody. *J. Virol.* **65**, 489–493.
14. Pinter, A., Honnen, W. J., Racho, M. E. & Tilley, S. A. (1993). A potent, neutralizing human monoclonal antibody against a unique epitope overlapping the CD4-binding site of HIV-1 gp120 that is broadly conserved across North American and African virus isolates. *AIDS Res. Hum. Retroviruses*, **9**, 985–996.
15. Thali, M., Olshevsky, U., Furman, C., Gabuzda, D., Posner, M. & Sodroski, J. (1991). Characterization of a discontinuous human immunodeficiency virus type 1 gp120 epitope recognized by a broadly reactive neutralizing human monoclonal antibody. *J. Virol.* **65**, 6188–6193.
16. Moore, J. P. & Sodroski, J. (1996). Antibody cross-competition analysis of the human immunodeficiency virus type 1 gp120 exterior envelope glycoprotein. *J. Virol.* **70**, 1863–1872.
17. Tilley, S. A., Honnen, W. J., Racho, M. E., Hilgartner, M. & Pinter, A. (1991). A human monoclonal antibody against the CD4-binding site of HIV1 gp120 exhibits potent, broadly neutralizing activity. *Res. Virol.* **142**, 247–259.
18. Thali, M., Furman, C., Ho, D. D., Robinson, J., Tilley, S., Pinter, A. & Sodroski, J. (1992). Discontinuous, conserved neutralization epitopes overlapping the CD4-binding region of human immunodeficiency virus type 1 gp120 envelope glycoprotein. *J. Virol.* **66**, 5635–5641.
19. Laal, S., Burda, S., Gorny, M. K., Karwowska, S., Buchbinder, A. & Zolla-Pazner, S. (1994). Synergistic neutralization of human immunodeficiency virus type 1 by combinations of human monoclonal antibodies. *J. Virol.* **68**, 4001–4008.
20. Zhang, M. Y., Xiao, X., Sidorov, I. A., Choudhry, V., Cham, F., Zhang, P. F. *et al*. (2004). Identification and characterization of a new cross-reactive human immunodeficiency virus type 1-neutralizing human monoclonal antibody. *J. Virol.* **78**, 9233–9242.
21. Zhang, M. Y., Shu, Y., Phogat, S., Xiao, X., Cham, F., Bouma, P. *et al*. (2003). Broadly cross-reactive HIV neutralizing human monoclonal antibody Fab selected by sequential antigen panning of a phage display library. *J. Immunol. Methods*, **283**, 17–25.
22. Saphire, E. O., Parren, P. W., Pantophlet, R., Zwick, M. B., Morris, G. M., Rudd, P. M. *et al*. (2001). Crystal structure of a neutralizing human IGG against HIV-1: a template for vaccine design. *Science*, **293**, 1155–1159.
23. D'Souza, M. P., Milman, G., Bradac, J. A., McPhee, D., Hanson, C. V. & Hendry, R. M. (1995). Neutralization of primary HIV-1 isolates by anti-envelope monoclonal antibodies. *AIDS*, **9**, 867–874.
24. D'Souza, M. P., Livnat, D., Bradac, J. A. & Bridges, S. H. (1997). Evaluation of monoclonal antibodies to human immunodeficiency virus type 1 primary isolates by neutralization assays: performance criteria for selecting candidate antibodies for clinical trials. AIDS Clinical Trials Group Antibody Selection Working Group. *J. Infect. Dis.* **175**, 1056–1062.
25. Kwong, P. D., Doyle, M. L., Casper, D. J., Cicala, C., Leavitt, S. A., Majeed, S. *et al*. (2002). HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. *Nature*, **420**, 678–682.
26. Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J. & Hendrickson, W. A. (1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*, **393**, 648–659.
27. Huang, C. C., Venturi, M., Majeed, S., Moore, M. J., Phogat, S., Zhang, M. Y. *et al*. (2004). Structural basis of tyrosine sulfation and VH-gene usage in antibodies that recognize the HIV type 1 coreceptor-binding site on gp120. *Proc. Natl Acad. Sci. USA*, **101**, 2706–2711.
28. Moulard, M., Phogat, S. K., Shu, Y., Labrijn, A. F., Xiao, X., Binley, J. M. *et al*. (2002). Broadly cross-reactive HIV-1-neutralizing human monoclonal Fab selected for binding to gp120-CD4-CCR5 complexes. *Proc. Natl Acad. Sci. USA*, **99**, 6913–6918.
29. Darbha, R., Phogat, S., Labrijn, A. F., Shu, Y., Gu, Y., Andrykovitch, M. *et al*. (2004). Crystal structure of the broadly cross-reactive HIV-1-neutralizing Fab X5 and fine mapping of its epitope. *Biochemistry*, **43**, 1410–1417.
30. Ofek, G., Tang, M., Sambor, A., Katinger, H., Mascola, J. R., Wyatt, R. & Kwong, P. D. (2004). Structure and mechanistic analysis of the anti-human immunodeficiency virus type 1 antibody 2F5 in complex with its gp41 epitope. *J. Virol.* **78**, 10724–10737.
31. Zwick, M. B., Komori, H. K., Stanfield, R. L., Church, S., Wang, M., Parren, P. W. *et al*. (2004). The long third complementarity-determining region of the heavy chain is important in the activity of the broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2F5. *J. Virol.* **78**, 3155–3161.
32. Cardoso, R. M., Zwick, M. B., Stanfield, R. L., Kunert, R., Binley, J. M., Katinger, H. *et al*. (2005). Broadly neutralizing anti-HIV antibody 4E10 recognizes a helical conformation of a highly conserved fusion-associated motif in gp41. *Immunity*, **22**, 163–173.
33. Choe, H., Li, W., Wright, P. L., Vasilieva, N., Venturi, M., Huang, C. C. *et al*. (2003). Tyrosine sulfation of human antibodies contributes to recognition of the CCR5 binding region of HIV-1 gp120. *Cell*, **114**, 161–170.
34. Xiang, S. H., Farzan, M., Si, Z., Madani, N., Wang, L., Rosenberg, E. *et al*. (2005). Functional mimicry of a human immunodeficiency virus type 1 coreceptor by a neutralizing monoclonal antibody. *J. Virol.* **79**, 6068–6077.
35. Wilkinson, R. A., Piscitelli, C., Teintze, M., Cavacini, L. A., Posner, M. R. & Lawrence, C. M. (2005). Structure of the Fab fragment of F105, a broadly

reactive anti-human immunodeficiency virus (HIV) antibody that recognizes the CD4 binding site of HIV type 1 gp120. *J. Virol.* **79**, 13060–13069.
36. Chen, B., Vogan, E. M., Gong, H., Skehel, J. J., Wiley, D. C. & Harrison, S. C. (2005). Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature*, **433**, 834–841.
37. Ramachandran, G. N. & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Advan. Protein Chem.* **23**, 283–438.
38. Milner-White, E., Ross, B. M., Ismail, R., Belhadj-Mostefa, K. & Poet, R. (1988). One type of gamma-turn, rather than the other gives rise to chain-reversal in proteins. *J. Mol. Biol.* **204**, 777–782.
39. Kontou, M., Leonidas, D. D., Vatzaki, E. H., Tsantili, P., Mamalaki, A., Oikonomakos, N. G. *et al*. (2000). The crystal structure of the Fab fragment of a rat monoclonal antibody against the main immunogenic region of the human muscle acetylcholine receptor. *Eur. J. Biochem.* **267**, 2389–2397.
40. Poulas, K., Eliopoulos, E., Vatzaki, E., Navaza, J., Kontou, M., Oikonomakos, N. *et al*. (2001). Crystal structure of Fab198, an efficient protector of the acetylcholine receptor against myasthenogenic antibodies. *Eur. J. Biochem.* **268**, 3685–3693.
41. Davies, D. R., Padlan, E. A. & Segal, D. M. (1975). Three-dimensional structure of immunoglobulins. *Annu. Rev. Biochem.* **44**, 639–667.
42. Davies, D. R. & Metzger, H. (1983). Structural basis of antibody function. *Annu. Rev. Immunol.* **1**, 87–117.
43. Amzel, L. M. & Poljak, R. J. (1979). Three-dimensional structure of immunoglobulins. *Annu. Rev. Biochem.* **48**, 961–997.
44. Novotny, J., Bruccoleri, R., Newell, J., Murphy, D., Haber, E. & Karplus, M. (1983). Molecular anatomy of the antibody binding site. *J. Biol. Chem.* **258**, 14433–14437.
45. Rothlisberger, D., Honegger, A. & Pluckthun, A. (2005). Domain interactions in the Fab fragment: a comparative evaluation of the single-chain Fv and Fab format engineered with variable domains of different stability. *J. Mol. Biol.* **347**, 773–789.
46. Vargas-Madrazo, E. & Paz-Garcia, E. (2003). An improved model of association for VH-VL immunoglobulin domains: asymmetries between VH and VL in the packing of some interface residues. *J. Mol. Recogn.* **16**, 113–120.
47. Chothia, C., Novotny, J., Bruccoleri, R. & Karplus, M. (1985). Domain association in immunoglobulin molecules. The packing of variable domains. *J. Mol. Biol.* **186**, 651–663.
48. Cohen, G. H., Sheriff, S. & Davies, D. R. (1996). Refined structure of the monoclonal antibody HyHEL-5 with its antigen hen egg-white lysozyme. *Acta Crystallog. sect. D*, **52**, 315–326.
49. Cohen, G. H., Silverton, E. W., Padlan, E. A., Dyda, F., Wibbenmeyer, J. A., Willson, R. C. & Davies, D. R. (2005). Water molecules in the antibody-antigen interface of the structure of the Fab HyHEL-5-lysozyme complex at 1.7 Å resolution: comparison with results from isothermal titration calorimetry. *Acta Crystallog. sect. D*, **61**, 628–633.
50. Schuermann, J. P., Henzl, M. T., Deutscher, S. L. & Tanner, J. J. (2004). Structure of an anti-DNA fab complexed with a non-DNA ligand provides insights into cross-reactivity and molecular mimicry. *Proteins: Struct. Funct. Genet.* **57**, 269–278.
51. Tormo, J., Stadler, E., Skern, T., Auer, H., Kanzler, O., Betzel, C. *et al*. (1992). Three-dimensional structure of the Fab fragment of a neutralizing antibody to human rhinovirus serotype 2. *Protein Sci.* **1**, 1154–1161.
52. Al-Lazikani, B., Lesk, A. M. & Chothia, C. (1997). Standard conformations for the canonical structures of immunoglobulins. *J. Mol. Biol.* **273**, 927–948.
53. Kabat, E. A., Wu, T. T., Perry, H. M., Gottesmann, K. S. & Foeller, C. (1991). *Sequences of Proteins of Immunological Interest*. NIH Publication no. 91-3242, US Department of Health and Human Services.
54. Tomlinson, I. M., Cox, J. P., Gherardi, E., Lesk, A. M. & Chothia, C. (1995). The structural repertoire of the human V kappa domain. *EMBO J.* **14**, 4628–4638.
55. Lefranc, M. P., Giudicelli, V., Kaas, Q., Duprat, E., Jabado-Michaloud, J., Scaviner, D. *et al*. (2005). IMGT, the international ImMunoGeneTics information system. *Nucl. Acids Res.* **33**, D593–D597.
56. Pruitt, K. D., Tatusova, T. & Maglott, D. R. (2005). NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucl. Acids Res.* **33**, D501–D504.
57. Morea, V., Lesk, A. M. & Tramontano, A. (2000). Antibody modeling: implications for engineering and design. *Methods*, **20**, 267–279.
58. Morea, V., Tramontano, A., Rustici, M., Chothia, C. & Lesk, A. M. (1998). Conformations of the third hypervariable region in the VH domain of immunoglobulins. *J. Mol. Biol.* **275**, 269–294.
59. Furukawa, K., Shirai, H., Azuma, T. & Nakamura, H. (2001). A role of the third complementarity-determining region in the affinity maturation of an antibody. *J. Biol. Chem.* **276**, 27622–27628.
60. Shirai, H., Kidera, A. & Nakamura, H. (1996). Structural classification of CDR-H3 in antibodies. *FEBS Letters*, **399**, 1–8.
61. Shirai, H., Kidera, A. & Nakamura, H. (1999). H3-rules: identification of CDR-H3 structures in antibodies. *FEBS Letters*, **455**, 188–197.
62. Wang, J. H., Yan, Y. W., Garrett, T. P., Liu, J. H., Rodgers, D. W., Garlick, R. L. *et al*. (1990). Atomic structure of a fragment of human CD4 containing two immunoglobulin-like domains. *Nature*, **348**, 411–418.
63. Kaur, H. & Raghava, G. P. (2004). Prediction of alpha-turns in proteins using PSI-BLAST profiles and secondary structure information. *Proteins: Struct. Funct. Genet.* **55**, 83–90.
64. MacCallum, R. M., Martin, A. C. & Thornton, J. M. (1996). Antibody-antigen interactions: contact analysis and binding site topography. *J. Mol. Biol.* **262**, 732–745.
65. Huang, C. C., Stricher, F., Martin, L., Decker, J. M., Majeed, S., Barthe, P. *et al*. (2005). Scorpion-toxin mimics of CD4 in complex with human immunodeficiency virus gp120 crystal structures, molecular mimicry, and neutralization breadth. *Structure (Camb)*, **13**, 755–768.
66. Janin, J. (2005). Assessing predictions of protein-protein interaction: the CAPRI experiment. *Protein Sci.* **14**, 278–283.
67. Lei, H. & Duan, Y. (2004). Incorporating intermolecular distance into protein-protein docking. *Protein Eng. Des. Sel.* **17**, 837–845.
68. Barbas, C. F., Burton, D. R., Scott, J. K. & Silverman, G. J. (2001). *Phage Display: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
69. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Macromol. Crystallog.* **276**, 307–326.

70. Navaza, J. (2001). Implementation of molecular replacement in AMoRe. *Acta Crystallog. sect. D*, **7**, 1367–1372.
71. Anderson, W. F., Cygler, M., Braun, R. P. & Lee, J. S. (1988). Antibodies to DNA. *Bioessays*, **8**, 69–74.
72. Cauerhff, A., Braden, B. C., Carvalho, J. G., Aparicio, R., Polikarpov, I., Leoni, J. & Goldbaum, F. A. (2000). Three-dimensional structure of the Fab from a human IgM cold agglutinin. *J. Immunol.* **165**, 6422–6428.
73. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
74. Brunger, A. T., Adams, P. D. & Rice, L. M. (1997). New applications of simulated annealing in X-ray crystallography and solution NMR. *Structure*, **5**, 325–336.
75. Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. (1993). Procheck—a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
76. Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* **8**(52-6), 29.
77. Carter, P., Lesk, V. I., Islam, S. A. & Sternberg, M. J. (2005). Protein-protein docking using 3D-Dock in rounds 3, 4, and 5 of CAPRI. *Proteins: Struct. Funct. Genet.* **60**, 281–288.
78. Esnouf, R. M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph. Modelling*, **15**, 132–134.

Edited by I. Wilson

(Received 5 August 2005; received in revised form 13 December 2005; accepted 15 December 2005)
Available online 9 January 2006